A Study on the Perception of Fundamental Frequency Declination by Japanese Speakers.

Shimizu, Katsumasa; Dantsuji, Masatake

STUDIA PHONOLOGICA XV (1981)

A Study on the Perception of Fundamental Frequency Declination by Japanese Speakers

Katsumasa SHIMIZU and Masatake DANTSUJI

0. Purpose

The aims of the present study are to examine the perception of fundamental frequency (Fo) contours by Japanese speakers, specifically to examine the normalization effect in perception for declination of Fo contours and its relationship with the length of stimuli. The effect in the perception of prosodic phenomena is one of the issues which have recently attracted the attention of a number of investigators in the related fields. Many attempts have been made on the causes of declination in production of Fo contours and its linguistic significance, but studies on normalization in perception are rather limited. In the present study, we shall examine how Japanese speakers perceive changes of Fo contours in relation with the length of stimuli and how they are sensitive to them in the perception of Fo declination and shall try to clarify the relation between production and perception of Fo contours.

1. Introduction

Fundamental frequency is one of the acoustic parameters which form prosodic patterns such as stress, tone, and intonation in a language and the acoustic correlate is perceptually manifested as pitch. It is also known that Fo contours carry linguistic information to phrase the stream of speech sounds and to perceive junctures in utterances. Studies on the nature of Fo contours in production and perception have been quite numerous; in production a number of experiments have been made to clarify physiological mechanism of vocal-folds vibration (Levitt & Rabiner, 1971; Collier, 1975; Sawashima & Cooper (ed.), 1977), aerodynamic measurements (Brown et al., 1970; Brown & McGlone, 1974) and its relationship with linguistic significance, while in perception there have been reports on the relationship between acoustic parameters and perception (Lieberman, 1965, 1977; Peck, 1969). Although it is pointed out that there are considerable variations in the production of Fo contours, it has recently been observed that Fo contours have a tendency to go downward continuously over the syntactically significant unit and breath length unit, and there are cases in some languages in Africa that high tone in the phrase-final position may be lower in physical scale than low tone

Katsumasa SHIMIZU (清水克正): Instructor, Dept. of Linguistics, Kyoto University, and also a faculty member of Nagoya Gakuin University.
Masatake DANTSUJI (壇辻正剛): Graduate Student, Dept. of Linguistics, Kyoto University. The authors are doing research under the direction of Dr. Tatsuo Nishida, Professor of Linguistics, Kyoto University.

in the phrase-initial position.[1] Downward movement of Fo contours is called declination and is said to have a linguistic function to indicate the presence of syntactically significant unit. Furthermore, degree of declination is affected by the length of the unit; the degree is less in longer syntactic unit, while it is greater in shorter one (Maeda, 1976). For the continuously declining Fo contours which may be found as a universal phenomenon, several physiological explanations have been proposed in terms of the lowering of larynx height, gradual reduction of subglottal pressure and "purposeful" changes of vocal-folds tension.[2] Although these may be attributed to the cause of declination, each explanation has its less convincing aspects and needs further elaborations.

In the perception of Fo contours, it has long been known that the perception of pitch levels does not necessarily synchronize with physical and acoustic scale.[3] It has also been known that, in relation with the declination effect in production, listeners tend to judge two syllables in one phrase to be equal in pitch, even if the second syllable is somewhat lower than the first in the acoustic scale. Listeners tend to normalize for declination in the perception of Fo contours, and we term this phenomenon perceptual normalization. In connection with this, Pierrehumbert (1979) carried out the experiments on the perception of Fo declination in terms of pitch range of the stimuli and amplitude contour, and mentioned the same results on the perception of two peaks in a nonsense sentence. We think it necessary to further elaborate the points made clear by Pierrehumbert (1979) and to examine how listeners are sensitive to the changes of Fo contours in initial and final positions and how the length of stimuli affects normalization in perception.

2. Experimental Methods

2.1 General experimental procedure

The stimuli for the experiments consist of three synthetic sets, depending on the number of syllables; three-syllable set /i o i/, five-syllable set /i e a o i/, and seven-syllable set /i e a o a o i/. The set /i e a o i/ means "(The) house is blue." in Japanese, if forced, while the sets of /i o i/ and /i e a o a o i/ do not make sense. For five- and seven-syllable sets, two series of stimuli were prepared for each set; concave type contour and convex type contour. Since each syllable is 250 msec in duration, including onset and offset, the whole duration is 750 msec for /i o i/, 1,250 msec for /i e a o i/ and 1,750 msec for

1. A gradual fall of pitch is exhibited in many languages and is called downdrift in African languages, whose process is considered to be automatic and non-morphological. The slope of downward movement of pitch is over a stretch of speech which Ohala (1978) termed as a "phonological phrase", a "breath group", or a "syntagm".
2. The causes of declination of pitch have been one of the controversial issues in phonetics. Ohala (1978) points out three hypotheses; the first is about the relation between larynx height and pitch, and the larynx progressively lowers during the breath group. The second is about a gradual reduction of subglottal pressure due to the pulmonic system. The third is that changes in tension of laryngeal muscles cause changes of pitch. Collier (1975) mentions that a gradual declination of pitch is controlled by subglottal pressure and the rapid changes of pitch are caused by the activity of cricothyroid muscle.
3. Lieberman (1965), pp. 52–53.

/i e a o a o i/. Formant frequencies of /i/, /e/, /a/, and /o/ are fixed at some specified frequencies, and Fo of final syllable /i/ in each set is a variable and is systematically varied at 9 equal steps.[4] All stimuli were prepared by the terminal analog speech synthesizer of the Department of Information Science, Kyoto University. The followings are the list of the stimuli for the experiments.

Set of Synthetic Stimuli

Stimulus Set	Contour Type	Fo of Initial /i/	Fo of Middle Syllable	Varying Range of Fo of Final /i/
/i o i/	Concave i o i	150 Hz	130 Hz	140–156 Hz
/i e a o i/	Concave i e a o i	150	130	140–156
	Convex i e a o i	130	150	120–136
/i e a o a o i/	Concave i e a o a o i	150	130	140–156
	Convex i e a o a o i	130	150	120–136

2.2 Experiment 1

This experiment was designed to examine the perception of Fo contours in three-syllable /i o i/. Fo values of the initial /i/ (hereinafter $/i/_1$) and the second /o/ were kept constant to 150 Hz and 130 Hz, respectively. 9 step stimuli were prepared with varying Fo of the final /i/ (hereinafter $/i/_2$) from 140 Hz to 156 Hz with 2 Hz equal step. Each syllable is 250 msec in duration, thus making the whole duration 750 msec. Each of /i o i/ stimuli was randomly presented 10 times, totalling 90 stimuli. Subjects were 30 university students who were phonetically naive, and were instructed to judge whether pitch of $/i/_2$ is higher, equal to, or lower than that of $/i/_1$. The results can be shown in Table 1 and Figure 1.

The results of Experiment 1 show that 78.7% of subjects judged that $/i/_2$ was equal to or higher than $/i/_1$ in the perception of pitch level when Fo value of $/i/_2$ was 146 Hz, 4 Hz lower than $/i/_1$. The percentage in which $/i/_2$ is judged equal to $/i/_1$ holds a major portion of distribution at Fo values of 146, 148 and 150 Hz. Furthermore, 27.0% of

4. All stimuli in the present experiments were synthesized at the Dept. of Information Science of Kyoto University. Acoustic details of each vowel can be shown as follows:

Formant	/i/	/e/	/a/	/o/	Bandwidth
F_5	4,500 Hz	4,500 Hz	4,500 Hz	4,500 Hz	281 Hz
F_4	3,500	3,500	3,500	3,500	175
F_3	3,000	2,550	2,600	2,600	200
F_2	2,100	2,100	1,200	880	130
F_1	300	510	700	470	80
Duration	250 msec	250	250	250	

Table 1. The results on judgement of relative pitch of $/i/_2$ in comparison with $/i/_1$ in the stimuli of / i | o | i /

Fo of $/i/_1$	Fo of $/i/_2$	Higher	Equal	Lower
150 Hz	140 Hz	5.3%	12.3%	82.4%
	142	6.3	20.0	73.7
	144	6.0	25.3	68.7
	146	14.7	64.0	21.3
	148	13.3	70.3	16.4
	150	27.0	70.0	3.0
	152	53.3	43.0	3.7
	154	78.0	19.7	2.3
	156	85.0	11.3	3.7

Higher: $/i/_2$ is judged higher than $/i/_1$ in pitch
Equal: $/i/_2$ is judged equal to $/i/_1$ in pitch
Lower: $/i/_2$ is judged lower than $/i/_1$ in pitch

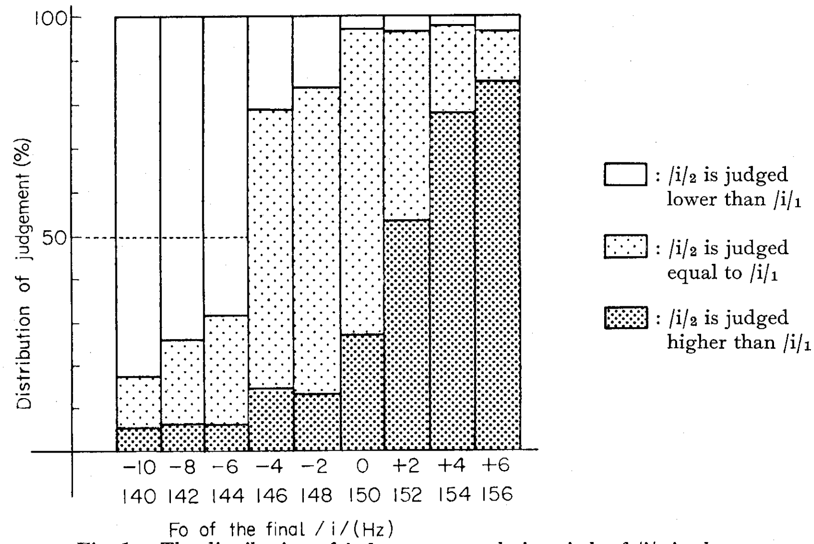

Fig. 1. The distribution of judgement on relative pitch of $/i/_2$ in three-syllable / i | o | i /

subjects judged that $/i/_2$ was higher than $/i/_1$ in pitch at 150 Hz which is acoustically equal to that of $/i/_1$. The results indicate that even if Fo of $/i/_2$ is 4 Hz lower than that of $/i/_1$, more than half of the subjects judged that both /i/s are the same in pitch. Taking 50% of equal and higher responses as a base to correct Fo declination, it can be assumed that listeners normalized for declination of 4 Hz in Fo contours in the three-syllable /i o i/.

2.3 Experiment 2

This experiment was designed to examine the perception of Fo contours of five-syllable stimuli /i e a o i/. The stimuli were synthesized in the same way as in Experiment 1. Two 9-step series were prepared; one series has concave type contour, keeping Fo of $/i/_1$ constant to 150 Hz and varying Fo of $/i/_2$ from 140 to 156 Hz with 2 Hz equal step. Fo values in middle syllables /e a o/ were kept constant to 130 Hz. The other series has convex type contour, keeping Fo of $/i/_1$ constant to 130 Hz and varying Fo of $/i/_2$ from 120 to 136 Hz with 2 Hz equal step. Each Fo value of middle syllables /e a o/ was 150 Hz. The stimuli in both series were 1,250 msec in whole duration. In both series, each of the /i e a o i/ stimuli was randomly presented 10 times, totalling 90 presentations of stimuli. Subjects were 40 university students for the perception of concave type series, and were 38 students for the perception of convex type series. In both series, subjects were instructed to judge the pitch height of $/i/_2$ in comparison with $/i/_1$ as in Experiment 1. The results for concave type can be shown in Table 2a and Figure 2a.

As shown in Table 2a and Figure 2a, 79.2% of subjects judged that $/i/_2$ was equal to or higher than $/i/_1$ in pitch when Fo of $/i/_2$ was 146 Hz. The distribution in Figure 2a shows that more than half of the subjects judged $/i/_2$ to be equal to $/i/_1$ when Fo value

Table 2a. The results on judgement of relative pitch of $/i/_2$ in comparison with $/i/_1$ in the stimuli of concave type /i|e a o|i/

Fo of $/i/_1$	Fo of $/i/_2$	Higher	Equal	Lower
150 Hz	140 Hz	4.0%	16.0%	80.0%
	142	7.0	22.5	70.5
	144	9.5	31.0	59.5
	146	17.7	61.5	20.8
	148	16.0	72.0	12.0
	150	22.0	69.0	9.0
	152	58.5	36.5	5.0
	154	76.4	17.8	5.8
	156	88.5	8.5	3.0

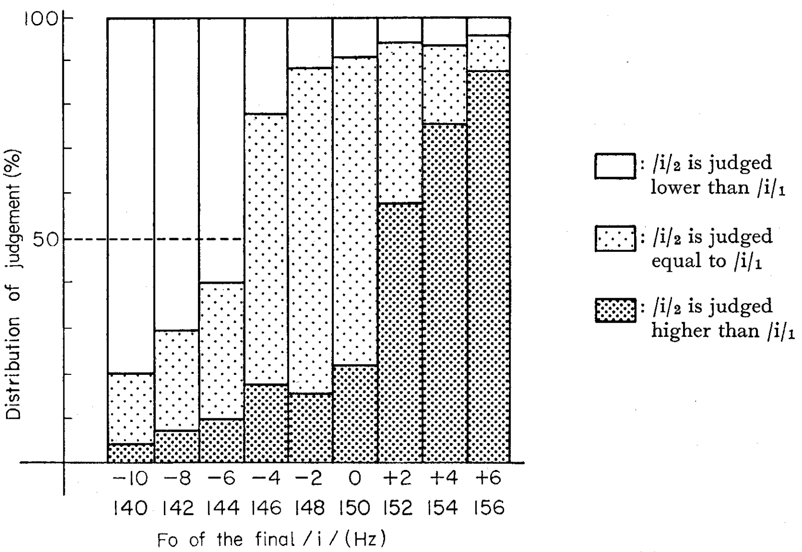

Fig. 2a. The distribution of judgement on relative picth of $/i/_2$ in five-syllable concave type /i|e a o|i/

was 2 to 4 Hz lower than that of /i/$_1$. There was a steplike increase for equal judgement at Fo value of 146 Hz and this indicates that subjects made a categorical judgement in the perception of two peaks of five-syllable stimuli. Taking 50% as a base as in Experiment 1, it can be assumed that listeners normalized for declination of 4 Hz in Fo contours in the concave type / i | e a o | i /.

Next, the results in the perception of convex type / i | e a o | i / can be shown in Table 2b and Figure 2b.

The results in Table 2b and Figure 2b show that 51.8% of subjects judged that /i/$_2$ was equal to or higher than /i/$_1$ in the perception of pitch when Fo of /i/$_2$ was 124 Hz, 6 Hz lower than that of /i/$_1$. It can be assumed that listeners normalized for declination of 6 Hz in convex type. The response data for equal judgement showed successive increments as Fo of /i/$_2$ was increased, and more than half of the subjects judged that /i/$_2$ was equal to /i/$_1$ when Fo of /i/$_2$ was 128, 130, and 132 Hz. Contrary

Table 2b. The results on judgement of relative pitch of /i/$_2$ in comparison with /i/$_1$ in the stimuli of convex type / i | e a o | i /

Fo of /i/$_1$	Fo of /i/$_2$	Higher	Equal	Lower
130 Hz	120 Hz	11. 1%	6. 6%	82. 3%
	122	13. 2	18. 2	68. 6
	124	19. 2	32. 6	48. 2
	126	18. 7	38. 4	42. 9
	128	16. 3	51. 6	32. 1
	130	18. 4	68. 7	12. 9
	132	27. 6	58. 2	14. 2
	134	50. 8	32. 4	16. 8
	136	57. 6	23. 9	18. 5

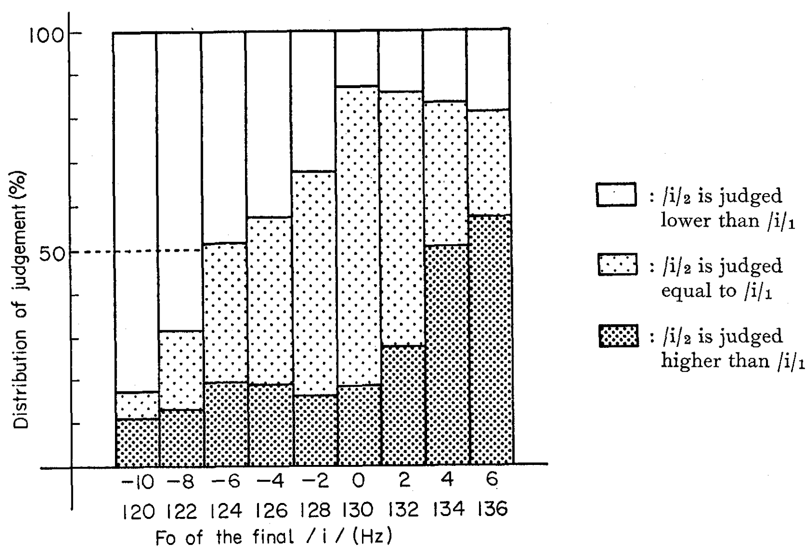

Fig. 2b. The distribution of judgement on relative pitch of /i/$_2$ in five-syllable convex type / i | e a o | i /

to the prediction, the number of subjects who judged /i/$_2$ lower than /i/$_1$ increased in the range of Fo of /i/$_2$ from 130 to 136 Hz.

The differences between concave type and convex type became evident in performance. There was a difference in normalization effect between two types; 4 Hz in concave type and 6 Hz in convex type, taking 50% of performance for equal and higher judgement as a base. The number of subjects who judged /i/$_2$ higher than /i/$_1$ was relatively small in convex type, compared with concave type, even at 132 Hz. This may be because subjects had a difficulty in judging pitch height of /i/$_2$ when Fo value of the preceding syllable was always higher in convex type. Furthermore, the performance for equal responses made a sharp and steplike increase at 146 Hz in concave type but not in convex type. In judging the relative height of Fo contours, it seems that subjects made a categorical-like decision in concave type.

2.4 Experiment 3

This experiment was designed to examine the perception of longer Fo contour of seven-syllable stimuli /i e a o a o i/. As in Experiment 2, two 9-step series were prepared; concave type and convex type contours. In the concave type stimuli, Fo of /i/$_1$ was 150 Hz and each Fo of the middle syllables /e a o a o/ was 130 Hz. Fo of /i/$_2$ was varied from 140 to 156 Hz with 2 Hz equal step. The whole duration of the stimuli was 1,750 msec. In the convex type stimuli, on the other hand, Fo of /i/$_1$ was 130 Hz and each of the middle syllables was 150 Hz. Fo of /i/$_2$ was varied from 120 to 136 Hz with 2 Hz equal step. In both series, each of the stimuli was randomly presented 10 times, totalling 90 stimuli. Subjects were 35 university students for the perception of concave type stimuli, while they were 33 students for the perception of convex type stimuli. They were naive in the perception test and were instructed to judge pitch height of /i/$_2$ in comparison with /i/$_1$. The results in the experiment of concave type /i e a o a o i/ can be shown in Table 3a and Figure 3a.

From the results in Table 3a and Figure 3a, it can be found that about half of the subjects judged that /i/$_2$ was equal to or higher than /i/$_1$ when Fo value of /i/$_2$ was 142 Hz, 8 Hz lower than /i/$_1$, and this performance went up to 94.8% at Fo value of

Table 3a. The results on judgement of relative pitch of /i/$_2$ in comparison with /i/$_1$ in the stimuli of concave type /i e a o a o i/

Fo of /i/$_1$	Fo of /i/$_2$	Higher	Equal	Lower
150 Hz	140 Hz	1.4%	7.1%	91.5%
	142	9.7	42.9	47.4
	144	10.3	54.0	35.7
	146	34.9	57.7	7.4
	148	37.7	56.9	5.4
	150	51.7	43.1	5.2
	152	66.6	29.1	4.3
	154	83.7	12.9	3.4
	156	90.6	8.6	0.8

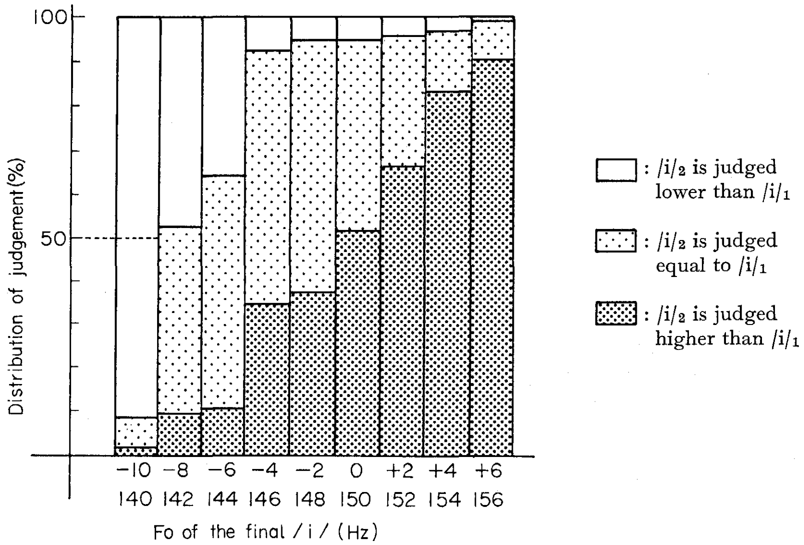

Fig. 3a. The distribution of judgement on relative pitch of $/i/_2$ in seven-syllable concave type / i | e a o a o | i /

150 Hz which is equal to that of $/i/_1$. As in Figure 2a, there was a sharp increase in the performance for equal judgement at 142 Hz. It can be assumed that listeners normalized for declination of 8 Hz in the perception of concave type / i | e a o a o | i /.

The results in the experiment of convex type / i | e a o a o | i / can be shown in Table 3b and Figure 3b.

The results in Table 3b and Figure 3b show that more than half of the subjects judged $/i/_2$ to be equal or higher than $/i/_1$ when Fo value of $/i/_2$ was 124 Hz. The performance for equal judgement exceeded 50% in the range from 124 to 132 Hz. It can be assumed that listeners normalized for declination of 6 Hz in the perception of convex type / i | e a o a o | i /. Furthermore, it can be noticed that the performance for higher judgement was extremely low even at Fo values of 132 and 134 Hz. The same remark as in the convex type of Experiment 2 may be made for this response data.

Table 3b. The results on judgement of relative pitch of $/i/_2$ in comparison with $/i/_1$ in the stimuli of convex type / i | e a o a o | i /

Fo of $/i/_1$	Fo of $/i/_2$	Higher	Equal	Lower
130 Hz	120 Hz	1.5%	5.3%	93.2%
	122	1.8	23.4	74.8
	124	4.2	51.3	44.5
	126	5.0	60.3	34.7
	128	5.6	63.8	30.6
	130	22.8	68.9	8.3
	132	42.1	52.2	5.7
	134	57.0	36.8	6.2
	136	76.0	17.5	6.5

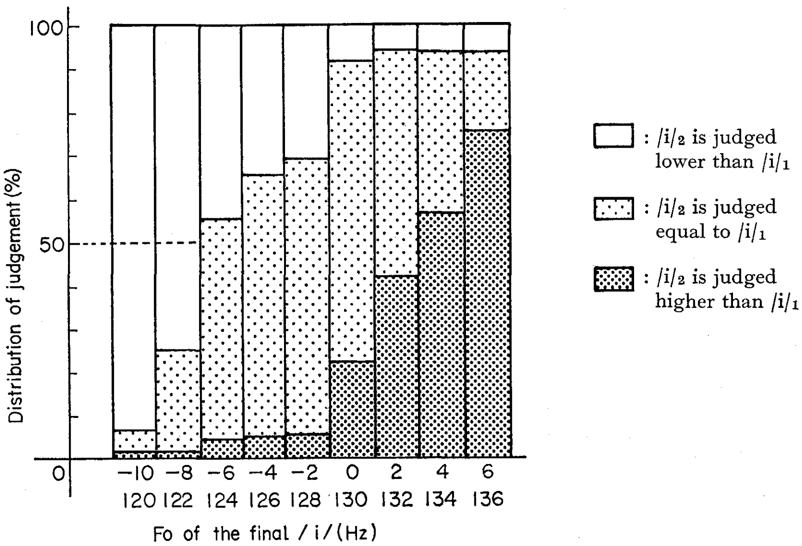

Fig. 3b. The distribution of judgement on relative pitch of $/i/_2$ in seven-syllable convex type /i e a o a o i/

When the utterance becomes longer as in seven syllables, it seems that subjects have a difficulty to judge whether $/i/_2$ is higher than $/i/_1$ in convex type.

3. Discussion

Five experiments were carried out in the present study to examine how speakers, specifically Japanese speakers, perceive changes of Fo contours and how normalization for pitch declination is correlated with such factors as the number of syllables in stimuli, Fo frequency in the initial syllable and pitch contour types. All stimuli are synthetic speech sounds consisting of vowels, and two syllables in the initial and final positions which were used for judgement of pitch were /i/. Fo frequency in the final position was systematically varied with 2 Hz equal step from 140 to 156 Hz for concave type and from 120 to 136 Hz for convex type. Subjects were instructed to judge whether $/i/_2$ with variable Fo frequency sounds higher, equal to or lower than $/i/_1$ in pitch.

3.1 Normalization effect in the perception of Fo contours

The results in the present experiments show that there exists normalization effect in perception of Fo contours. More than half of the subjects in each experiment judged that $/i/_2$ is equal to $/i/_1$, even if Fo of $/i/_2$ is 4 to 8 Hz lower than that of $/i/_1$. That is, two /i/s sound equal in pitch, even if $/i/_2$ is lower than $/i/_1$ in acoustic scale, and psychological judgement in the perception of Fo contours is not directly correlated with acoustic and physical properties of the stimuli. Normalization in the perception of pitch is positively related with the number of syllables in stimuli. The experiments were carried out for three-, five- and seven-syllable length stimuli. It is reasonable to suppose that the number of intervening syllables affects the perception of pitch of two /i/s and normalization is less in smaller number of syllables than in greater number of syllables because

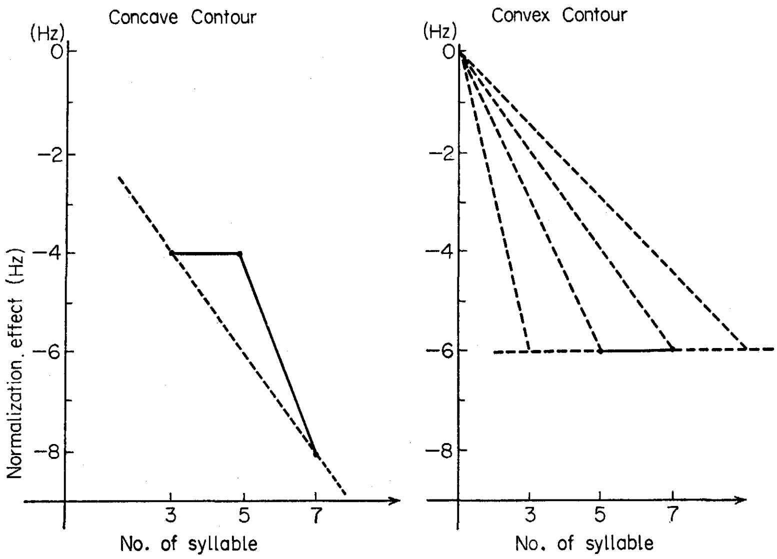

Fig. 4. The relationship between normalization effect and the number of syllables in stimuli

of recency effect in memory load. The results show that as the number of syllables increases, there is a tendency, though not complete, to have a greater effect. In the concave type contours, this tendency was more clearly manifested; 4 Hz for three-syllable set, 4 Hz for five-syllable set and 8 Hz for seven-syllable set. In the convex type contours, however, there was no noticeable difference between five- and seven-syllable sets; 6 Hz of normalization effect was observed in both sets. The relationship between the effect and the number of syllable in stimuli can be shown in Figure 4.

As seen in Figure 4, there was a difference in normalization between concave type and convex type contours. In the concave type, subjects were asked to judge the pitch of two *peak* syllables, and Fo frequency of the syllable preceding $/i/_2$ was always lower than $/i/_2$. On the other hand, in the convex type, subjects were asked to judge the pitch of two *bottom* syllables and Fo frequency of the syllable preceding $/i/_2$ was always higher than $/i/_2$. The results indicate that Fo frequency of the preceding syllable affects auditory correction in the perception of pitch, and when it is higher than $/i/_2$, subjects have a difficulty to compare $/i/_2$ with $/i/_1$ in pitch. This seems to cause the difference in the range of normalization effect between two types of contours.

In connection with the above, it is worthwhile to examine normalization effect from the point of view of Fo frequency of $/i/_1$ and the response for equal judgement. Fo frequency of the initial $/i/_1$ was set at 150 Hz for concave type and at 130 Hz for convex type. Based on the response data in which subjects judged two /i/s to be equal in pitch, excluding the data for higher or lower judgement, we examined Fo frequency of $/i/_2$, at which highest performance was observed for equal judgement. The results can be shown in Table 4.

For concave type contours, it can be found that a greater normalization effect was observed in seven-syllable set than in three- and five-syllable sets. The effect was 3.1 Hz

Table 4 The results on predicted Fo frequency for equal judgement

	No. of syllable	Fo of $/i/_1$	Fo of $/i/_2$ judged most appropriate for equal judgement	Predicted Fo of $/i/_2$ for equal judgement
Concave Contour	3	150 Hz	148 Hz (−2)	148.2 Hz (s.d. 3.7)
	5	150	148 (−2)	147.8 (3.7)
	7	150	146 (−4)	146.9 (3.8)
Convex Contour	5	130	130 (0)	129.2 (4.0)
	7	130	130 (0)	128.5 (3.9)

at a 0.01 level of significance in seven-syllable set, while it was 1.8 Hz and 2.2 Hz in three- and five-syllable sets, respectively. The difference between three- and five syllable sets was not so significant. For convex type contours, on the other hand, the normalization effect was 0.8 Hz and 1.5 Hz in five- and seven-syllable sets, respectively, at a 0.01 level of significance. It can be noticed that the effect in the convex type was less than in concave type, and, if taken only the data for equal judgement into account, the effect is not proportionally related with the number of syllables in stimuli.

Through a series of experiments on the perception of Fo contours, it has become evident that listeners normalize for declination of pitch more manifestly in concave type than in convex type. Listeners tend to judge that the second peak is equal to the first peak in pitch, even if Fo of the former is 4 Hz lower than that of the latter. This is in agreement with previous findings in the studies on perception of prosodic phenomena.[5]

Furthermore, a difference between concave type and convex type can be seen in the perceptual mode for equal judgement. Examining the distribution for equal judgement in Figures 1, 2a and 3a, there is an apparent categorical shift in decision, and this shift took place at −4 Hz in Fig. 1, −4 Hz in Fig. 2a, and −8 Hz in Fig. 3a. On the other hand, this shift is apparently absent in Fig. 2b and in Fig. 3b for convex type. In concave type, listeners tend to normalize for declination in a categorical manner. This implies that they tend to normalize for declination within a limited range of Fo difference between $/i/_2$ and $/i/_1$. This perceptual mode is more clearly manifested in concave type than in convex type.

3.2 Perceptual compensation for pitch declination

As mentioned earlier, it is generally known that pitch contours go gradually downward, and its physiological correlates for declination in production have been extensively studied in terms of the effects of laryngeal muscle activity, subglottal pressure and the position of larynx. Hakoda and Sato (1976) pointed out that a difference between the sentence initial Fo frequency and the final Fo frequency is more than 30 Hz in most cases of utterances, and the difference in production is greater in the case of female speakers. The normalization effect in perception works in a direction to compensate

5. Breckenridge (1977) and Pierrehumbert (1979) examined the auditory effect of pitch declination, and mentioned that listeners judged the second peak of pitch to be equal to the first peak even though the second was a few hertz lower than the first.

a gradual downward running movement of pitch in production. Since the effect is in the range from 4 to 6 Hz in perception, the effect to compensate the declination in production seems to be less influential. However, it is considered that pitch perception operates by intervals rather than absolute frequencies, and the normalization effect in perception which works in the upward direction is considered to have a linguistically significant role in the perception of Fo contours.[6]

In the production of Fo contours, it is also known that the slope is inversely related to the length of utterance. This inverse relationship is not found in the perception of pitch; rather normalization seems to be proportional to the length of stimuli in a general trend, especially in concave type contours. The effect of predicted normalization was greater in seven-syllable set than in three- and five-syllable sets, but the difference between three- and five-syllable sets was not clearly manifested. It seems that there is some limited range for normalization in perception because of the memory load.

The results in the present experiments indicate that there is a considerable difference between concave and convex type contours; normalization in perception is greater in concave type than in convex type, and the categorical mode of perception is found in concave type. In carrying out experiments for convex type, subjects showed a difficulty in comparing pitch of two syllables separated by intervening syllables. This suggests that Fo frequency of the syllable preceding $/i/_2$ give an influence on the judgement of pitch, and if it is higher than that of $/i/_2$, a lesser normalization can be observed.

4. Summary

The present series of experiments increase our understanding on the perception of Fo contours. We have presented data here that speakers of Japanese normalize for declination in judging pitch height of synthetic stimuli and they judge that the final syllable is equal to the initial one in pitch even if Fo of the former is 4 to 8 Hz lower than that of the latter. This effect is more clearly demonstrated in concave type contours than in convex type contours. In connection with this, it can be argued that Fo of the syllable preceding the final one affects the normalization effect, and if the Fo is higher than that of the final one, the effect is weakened. Subjects tend to normalize for declination of Fo contours in a categorical manner in concave type and this implies that normalization works within a limited range of Fo difference between two peaks. Furthermore, normalization is proportional, though not consistent, to the length of stimuli, especially in concave type. Based on the response data for equal judgement, the effect was greater in seven-syllable set than in three-and five-syllable sets, but there was no noticeable difference between three- and five-syllable sets. Normalization in the perception of Fo contours can be accounted for by perceptual compensation for declination of pitch contours in production. The effect works in the direction to correct declination of Fo contours in production, and this direction is more perceptually significant than the absolute frequencies in exerting the effect.

6. Lehiste (1976), p. 230.

ACKNOWLEDGEMENTS

We would like to thank Dr. Tatsuo Nishida of Kyoto University who has assisted our research in linguistics in many ways. We also want to thank Prof. Toshiyuki Sakai, the Dept. of Information Science, Kyoto University, for providing their computer facilities and assisting to prepare synthetic stimuli, and Dr. Ilse Lehiste, The Ohio State University, for giving comments and suggestions on this paper.

REFERENCES

Breckenridge, J. (1977), "The declination effect", *J. Acoust. Soc. Am.*, Vol. 60, S90.

Brown, W. S. and McGlone, R. E. (1974), "Aerodynamic and acoustic study of stress in sentence productions", *J. Acoust. Soc. Am.*, Vol. 56, No. 3, pp. 971–974.

Brown, W. S., McGlone, R. E., Tarlow, A. and Shipp, T. (1970), "Intraoral air pressures associated with specific phonetic positions", *Phonetica*, 22, pp. 202–212.

Collier, R. (1975), "Physiological correlates of intonational patterns", *J. Acoust. Soc. Am.*, Vol. 58, No. 1, pp. 249–255.

Hakoda, K. and Sato, H. (1976), "Some characteristics of pitch contour in sentence speech", *J. Acoust. Soc. Japan*, pp. 109–110. (in Japanese)

't Hart, J. and Collier, R. (1979), "On the interaction of accentuation and intonation in Dutch", *Proceedings of the Ninth International Congress of Phonetic Sciences*, Vol. II, pp. 395–402.

Lehiste, I. (1976). "Suprasegmental features of speech", *Contemporary Issues in Experimental Phonetics* ed. by N. J. Lass, Academic Press, pp. 225–239.

Levitt, H. and Rabiner, L. R. (1971), "Analysis of fundamental frequency contours in speech", *J. Acoust. Soc. Am.*, Vol. 49, No. 2, pp. 569–582.

Lieberman, P. (1965), "On the acoustic basis of the perception of intonation by linguists", *Word*, Vol. 21, No. 1, pp. 40–54.

Lieberman, P. (1977), *Speech Physiology and Acoustic Phonetics*, Macmillan Publishing Co., Inc.

Maeda, S. (1976), *A Characterization of American English Intonation*, doctoral dissertation, MIT, Cambridge.

Nakatani, L. H. and Schaffer, J. A. (1978), "Hearing "words" without words: prosodic cues for word perception", *J. Acoust. Soc. Am.*, Vol. 63, No. 1, pp. 234–245.

Ohala, J. J. (1978), "Production of tone", *Tone: A Linguistic Survey*, ed. by V. A. Fromkin, Academic Press, pp. 5–39.

Olive, J. P. (1975), "Fundamental frequency rules for the synthesis of simple declarative sentences", *J. Acoust. Soc. Am.*, Vol. 57, No. 2, pp. 476–482.

Peck, C. W. (1969), *An Acoustic Investigation of the Intonation of American English*, Natural Language Studies No. 1, unpublished, Phonetics Laboratory, The Univ. of Michigan.

Pierrehumbert, J. (1979), "The perception of fundamental frequency declination", *J. Acoust. Soc. Am.*, Vol. 66, No. 2, pp. 363–369.

Pierrehumbert, J. (1980), *The Phonology and Phonetics of English Intonation*, doctoral dissertation, MIT, Cambridge.

Sawashima, M. and Cooper, F. S. ed. (1977), *Dynamic Aspects of Speech Production*, University of Tokyo Press.

Shimizu, K. and Dantsuji, M. (1980), "A study on perception of internal juncture in Japanese", *Studia Phonologica*, XIV, pp. 1–15.

Thorsen, N. (1978), "An acoustical investigation of Danish intonation", *J. of Phonetics*, Vol. 6, pp. 151–175.

Thorsen, N. (1980), "A study of the perception of sentence intonation—Evidence from Danish", *J. Acoust. Soc. Am.*, Vol. 67, No. 3, pp. 1014–1030.

Zee, E. (1978), "Duration and intensity as correlates of Fo", *J. of Phonetics*, Vol. 6, pp. 213–220.

(Aug. 31, 1981, received)